Слайд 1

**Слайд 2**

Дана робота містить огляд алгоритмів машинного навчання для виявлення аномальної поведінки об'єктів за допомогою камер з інтелектуальною аналітикою у сучасних системах відеоспостереження а також порівняльний аналіз ефективності алгоритмів у виявленні аномальної поведінки людини.

У сучасному світі безпека стала одним з головних пріоритетів, а системи відеоспостереження відіграють важливу роль у забезпеченні безпеки різних об’єктів.

Останнім часом відеоаналітика на основі штучного інтелекту набуває все більшої популярності та стає ключовим інструментом для виявлення аномалій у комплексних системах безпеки. В даному дослідженні ми зосередимося на аналізі сучасних систем відеоспостереження та методів обробки відеоінформації, а також на огляді методів детектування аномалій. Вивчення цих методів допоможе розробити нові та ефективні підходи для виявлення аномалій у реальному часі та підвищення ефективності комплексних систем безпеки [1].

**Слайд 3**

Системи відеоспостереження є невід'ємною частиною сучасних комплексних систем безпеки. Вони використовуються для моніторингу різних об'єктів та територій з метою забезпечення безпеки людей, майна, інфраструктури та інформаційних ресурсів. В рамках комплексних систем безпеки, системи відеоспостереження вирішують ряд ключових завдань.

**Слайд 4**

Однією з ключових проблем у забезпеченні безпеки даних з камер спостереження є виявлення аномальної поведінки. Аномалії можуть приймати різні форми, від спроб несанкціонованого доступу до підозрілих моделей руху в полі зору камери. Своєчасне і ефективне виявлення і реагування на ці аномалії має вирішальне значення для підтримки цілісності системи відеоспостереження і захисту людей і майна, що знаходяться під її наглядом.

Виявлення аномалій у даних камер спостереження є складним завданням через велику кількість даних, що генеруються, високу мінливість вмісту сцени та освітлення, а також потенційну можливість хибних спрацьовувань. Тому існує потреба в ефективній системі, яка може точно і ефективно виявляти аномалії в даних камер спостереження і попереджати персонал служби безпеки в режимі реального часу.

Застосування алгоритмів машинного навчання для аналізу відеопотоку дозволяє виявляти аномальну поведінку та потенційні загрози, навіть якщо вони відрізняються від типових сценаріїв, завдяки адаптивності та здатності до самонавчання та дозволяє системам відеоспостереження розпізнавати осіб, транспортні засоби, номерні знаки та інші об'єкти з високою точністю та швидкістю.

**Слайд 5**

Хочу представити найбільш найбільш ефективні алгоритми машинного навчання а також їхні переваги та недоліки.

Наступні алгоритми ML були обрані як найбільш придатні для порівняння їх ефективності в подальших дослідженнях в контексті виявлення аномалій в поведінці людини.

1) CNN (Convolutional Neural Network) – CNN широко використовуються для задач виявлення аномалій на основі відео завдяки своїй здатності фіксувати просторові відносини та витягувати локальні особливості із зображень та відеоданих.

2) LSTM (Long Short-Term Memory) – LSTM часто використовуються в задачах відеоаналізу через їхню здатність фіксувати довгострокові залежності в часових даних. Це особливо корисно при аналізі відео з камер спостереження для виявлення аномалій.

3) SVM (Support Vector Machine) – SVM є популярними моделями керованого навчання, які можна використовувати для виявлення аномалій шляхом максимізації різниці між різними класами в просторі ознак. SVM застосовуються в різних додатках, включаючи виявлення аномалій у відеоспостереженні, за допомогою функцій ядра для обробки нелінійно розділених даних.

4) RNN (Recurrent Neural Network) – RNN є класом нейронних мереж, які спеціально розроблені для обробки послідовних даних, що робить їх придатними для завдань відеоаналізу. Вони можуть вловлювати часові залежності в даних і були використані в різних додатках для виявлення аномалій у відеоспостереженні.

Ці алгоритми були обрані на основі їхньої поширеності в літературі та придатності для обробки відеоданих.

**Слайд 6**

Усі набори даних, що були розглянуті в цьому розділі, мають мітки для машинного навчання, і багато з них містять немічені відео для тестування алгоритмів на додаток до відеоданих, тому вибір найбільш підходящого набору даних базується на наступних факторах.

Набори даних, що містять відеозаписи поведінки людей у різних умовах, наприклад, відео з камер спостереження, найкраще підходять для навчання алгоритмів машинного навчання для виявлення аномалій у поведінці людей. Ці набори даних повинні містити достатньо аномалій, щоб алгоритм міг точно виявляти рідкісні події, і в той же час мати достатньо велику кількість нормальної поведінки, щоб забезпечити повне розуміння очікуваних дій. Крім того, набори даних, які є різноманітними з точки зору умов освітлення, кутів нахилу камери та типів аномалій, допоможуть забезпечити здатність алгоритму узагальнювати реальні сценарії. Набір даних, який відповідає цим критеріям – це UCF-Crime.

**Слайд 7**

Було обрано загальнодоступний набір даних UCF-crime, який містить відеозаписи різних аномальних і нормальних форм людської поведінки. База даних UCF-Crime – це масштабний набір даних із 128 годин відеозаписів. Він складається з 1900 довгих і не обрізаних відеозаписів реального спостереження з 13 реалістичними аномаліями, включаючи насильство, арешт, підпал, напад, дорожньо-транспортну пригоду, крадіжку, вибух, бійку, пограбування, стрілянину, крадіжку в магазині та вандалізм.

Етапи попередньої обробки включають завантаження набору даних і вилучення фіксованої кількості кадрів з кожного відео. Потім ці кадри попередньо обробляються шляхом перетворення їх у відтінки сірого та зміни розміру до фіксованої форми (64x64 пікселів). Результуючий набір даних складається зі списку відеокадрів (X) і відповідних міток, де 1 позначає нормальну поведінку, а 0 – аномальну.

Після завантаження та попередньої обробки набір даних розбивається на навчальний та тестовий набори у співвідношенні 80/20, де 80% даних використовується для навчання, а решта 20% – для тестування. Набір даних додатково обробляється, щоб зробити його придатним для різних алгоритмів машинного навчання. Для моделей CNN, LSTM та RNN значення пікселів кадрів нормалізуються шляхом ділення їх на 255, перетворюючи значення в діапазон [0, 1]. Для SVM-моделі кадри перетворюються в одновимірний формат і масштабуються за допомогою StandardScaler, щоб забезпечити однаковий масштаб ознак.

За допомогою цих кроків попередньої обробки набір даних перетворюється у формат, придатний для навчання та тестування обраних алгоритмів машинного навчання.

**Слайд 8 CNN**

Методологія використання CNN у цьому дослідженні включає кроки, що описані нижче [5]. Фреймворк CNN складається з двох основних компонентів. По-перше, механізм згортки ідентифікує та розрізняє різні особливості зображення для аналізу за допомогою процедури, відомої як вилучення особливостей. Цей процес включає в себе безліч комбінацій згортки або об'єднання шарів. По-друге, щільно з'єднаний шар приймає результати етапу згортки і прогнозує категорію зображення на основі попередньо вилучених ознак [7].

Метою підходу до вилучення ознак на основі CNN є зменшення кількості ознак у наборі даних при одночасній генерації нових, узагальнених ознак з оригінальної колекції. Схему архітектури CNN зображено на рисунку 3.1, який ілюструє різні рівні, що здіяні в цьому процесі.

**Слайд 9**

LSTM працює як з довготривалою пам'яттю LTM (Long Term Memory), так і з короткотривалою пам'яттю STM (Short Term Memory), використовуючи концепцію вентилів для спрощення і прискорення обчислень, як показано на рисунку 3.2.

1) Ворота забуття (Forget Gate) – отримує LTM і відкидає несуттєву інформацію.

2) Ворота вивчення (Learn Gate) – поточний вхід (подія) і STM об'єднуються, що дозволяє застосувати останні знання з STM до поточного входу.

3) Ворота запам'ятовування (Remember Gate) – об'єднує збережену інформацію LTM з STM і подією для створення оновленого LTM.

4) Використання воріт (Use Gate) – LTM, STM і подія використовуються цими воротами для прогнозування результатів поточної події, в результаті чого створюється оновлена STM.

**Слайд 10**

1) Вхідні дані. Вхідні дані для SVM включають попередньо оброблені ознаки, витягнуті з наборів даних. Кожна точка даних на вході представлена у вигляді вектора ознак у багатовимірному просторі ознак.

2) Функція ядра. Функція ядра відіграє вирішальну роль в архітектурі SVM. Це математична функція, яка відображає точки вхідних даних у простір вищої розмірності, що полегшує пошук розділяючої гіперплощини, навіть якщо дані не піддаються лінійному розділенню у вихідному просторі. Деякі поширені функції ядра включають лінійну, поліноміальну, радіально-базисну функцію RBF та сигмоїдну.

3) Опорні вектори. Вектори підтримки – це точки даних, які лежать найближче до межі рішення або гіперплощини. Вони сприяють визначенню положення оптимальної гіперплощини і є критично важливими для визначення межі між різними класами.

4) Маржа. Маржа – це відстань між розділювальною гіперплощиною та найближчими точками даних з кожного класу (опорними векторами). SVM намагається максимізувати цю відстань, щоб створити надійну границю рішення, яка допомагає мінімізувати помилку класифікації.

5) Кордон рішення або гіперплощина. Межа рішення або гіперплощина - це підпростір (лінія, площина або гіперплощина, залежно від розмірності вхідних даних), який розділяє точки даних на різні класи. Оптимальна гіперплощина - це та, яка найкраще розділяє класи, максимізуючи при цьому маржу.

6) Регуляризація. Регуляризація – це метод, який використовується в SVM для контролю компромісу між максимізацією маржі та мінімізацією помилки класифікації. Вона допомагає запобігти надмірному пристосуванню, вводячи штрафний член в задачу оптимізації. Параметр регуляризації, який часто позначають як «C», визначає баланс між максимізацією маржі та мінімізацією помилки класифікації [9].

**Слайд 11**

У рекурентній нейронній мережі RNN кожен вхідний елемент пов'язаний між собою. Спочатку RNN обробляє перший вхід, X(0), і генерує вихід, h(0). Згодом h(0) і X(1) об'єднуються як вхідні дані для наступного кроку. Таким же чином, h(1) і X(2) використовуються як вхідні дані для наступного кроку, і так далі. Такий підхід гарантує, що RNN зберігає контекстну інформацію протягом усього процесу навчання, як показано на рисунку 3.4.

**Слайд 12**

ROC-AUC є ефективною метрикою для порівняння продуктивності різних алгоритмів ML, оскільки вона враховує компроміс між TPR і FPR на різних порогових рівнях, надаючи єдине значення для для оцінки загальної продуктивності. Вище значення ROC-AUC вказує на кращу ефективність класифікації.

EER корисний для порівняння алгоритмів ML, оскільки він надає єдине скалярне значення, яке представляє оптимальний компроміс між хибнопозитивними та хибнонегативними результатами [20]. Нижчий EER вказує на кращу ефективність класифікації. Порівнюючи значення EER протестованих алгоритмів можна визначити, який алгоритм має кращий баланс між помилковими спрацьовуваннями і помилковими негативами, що робить його більш ефективним у виявленні аномальної поведінки людини.

**Слайд 13**

Модель CNN досягла показника AUC 0,8346, що свідчить про високу ефективність класифікації. Значення AUC ближче до 1 означає ідеальний класифікатор, тоді Таким чином, отримане значення AUC 0,8346 демонструє здатність моделі ідентифікувати аномалії зі значною точністю. З точки зору метрики EER, модель CNN дала значення 0,2439, що свідчить про те, що при оптимальному порозі класифікації. Нижче значення EER вказує на кращий баланс між хибними спрацьовуваннями та хибними неспрацьовуваннями, що призводить до покращення ефективності класифікації.

Модель LSTM у свою чергу продемонструвала ROC-AUC 0,7797 та EER 0,3.Таким чином, мережі LSTM демонструють відносно ефективну роботу у виявленні аномальної поведінки людини за допомогою даних з камер спостереження. Хоча результати не є оптимальними, підхід LSTM залишається життєздатним варіантом для задач виявлення аномалій.

**Слайд 14**

Показник ROC-AUC для алгоритму SVM склав 0,7281 а EER – 0,3780 що вказує на помірний рівень продуктивності в розрізненні аномальних і нормальних подій.

Цей результат демонструє потенціал алгоритму SVM для задач класифікації відео, але також підкреслює необхідність подальшої оптимізації та вивчення інших методів машинного навчання для підвищення ефективності класифікації.

Ефективність RNN-моделі ROC-AUC, 0,7602. Цей результат показує, що RNN-модель перевершила алгоритм SVM, але алгоритми LSTM і CNN все ще перевершують його в розрізненні аномальних і нормальних подій. Що стосується метрики EER, то модель RNN досягла значення 0,3292. Загалом, продуктивність моделі RNN демонструє досить добру роботу в задачі класифікації відео.

**Слайд 15**

На основі переваг і обмежень кожного алгоритму машинного навчання, можна зробити висновок, що CNN та LSTM можуть бути найбільш ефективними для виявлення аномалій поведінки, у тому числі таких, як виявлення інсайдерів, та аномальних ситуацій, як пожежа, задимлення, проникнення у приміщення сторонніх осіб у неробочий час, пронесення на контрольовану територію підозрілих предметів, тощо. Це пояснюється тим, що CNN володіє високою точністю в роботі з просторовими об'єктами, а LSTM здатний фіксувати часові залежності і підходить для часових рядів даних.

Слайд 16

Висновки можна побачити на слайді